



r語言

- 「=」 和 「←」 表示函數等於...
- 儲存成**character**的型態，要加上""
- (-4):(-2)，會呈現出-4,-3,-2
- 多個數字/字母時，如x=c(3,4,5)
 - c()組合括號內的元素成為一個清單或向量
 - 如：newObject <- c(5,12,36)，newObject是一個三維的向量
 - 要叫出其中一個時，x[2]，表示要叫出x序列中的第2個，也就會叫出4
 - 不要叫出其中一個時，x[-2]，就會顯示3 5
 - 叫出 大於小於等於某數字 的數字時，x[x>4]，就會出現5
- class(data1)，可以知道data1的資料類型
- elementOne ← **as.numeric(elementOne)**，把elementOne的資料類型轉為數字
- lengNew<-**length(newObject)**，lengNew中會有newObject內的資料數量
- 「ls()」可以看見現在有的函數（或從workspace也可查看）；
「rm(y)」可以消除y或其他函數
- seq(from=1, to=5, by=1)，一個從1到5，間隔1的數列，也可以直接寫成
seq(1,5,1)

```
> seq(from=1, to=5, by=1)
[1] 1 2 3 4 5
```

- rep(1, 4)，重複四次1

```
> rep(1, 4)
[1] 1 1 1 1
```

- `matrix(c(1,2,3,4,5,6), nrow=3, byrow=TRUE)`，代表有3列，`byrow=TRUE`代表一行一行開始排

```
> matrix(c(1,2,3,4,5,6), nrow=3, byrow=TRUE)
```

```
      [,1] [,2]
[1,]     1     2
[2,]     3     4
[3,]     5     6
```

```
> matrix(c(1,2,3,4,5,6), nrow=3, byrow=FALSE)
```

```
      [,1] [,2]
[1,]     1     4
[2,]     2     5
[3,]     3     6
```

- `matrixSub <- matrixOne[7:8,3:5]`，擷取`matrixOne`內的第7-8列和第3-5行
- `getwd()`查詢現在在用的文件檔案，`setwd()`設定要用的文件檔案，`setwd("/Users/sydney/Document/決策分析")`或`setwd("~/Document/決策分析")`
- `data1<-read.csv(file.choose(),header=T)`，用`r`選取已有的csv檔，`header=T`用來表示表頭是文字
- `dim(data1)`，可以知道`data1`有多少行列；
`head(data1)`、`tail(data1)`分別可以知道前後六列的資料

```
> dim(chapter09data)
[1] 5 7
> head(chapter09data)
# A tibble: 5 × 7
  ...1    ...2    ...3    ...4    ...5    ...6    ...7
  <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 壓力      1      2      3      4      5      6
2 8500    90.3   89.2   98.2   93.9   87.4   97.9
3 8700    92.5   89.5   90.6   94.7   87      95.8
4 8900    85.5   90.8   89.6   86.2   88      93.4
5 9100    82.5   89.5   85.6   87.4   78.9   90.7
```

- `mean(chapter09data$壓力)`，可以算出chapter09data中在壓力這一行的數值平均
另一個方法呈現如下，如果不要attach數據了，就`detach(chapter09data)`

```
> attach(chapter09data)
> mean(壓力)
[1] 8800
```

- `levels(gender)`，可以知道gender這一行裡的factor分別是什麼；
`mean(age[gender=="male"])`，可以計算出男性的平均年齡；
`maleover15 <- data1[gender=="male" & age>15,]`
- `summary(data1)`，可以知道data1每一行的數據分析
- 可以知道整個資料裡第1-3列「年齡大於10歲」的有哪些

```
> temp<-Age>10
> temp[1:3]
[1] FALSE  TRUE  TRUE
> temp1<-as.numeric(Age>13)
> temp1[1:3]
[1] 0 1 1
```

- `data2 ← cbind(data,data1)`, data2會是data和data1的結合
- `list[[1]]`, list中的第一個物件
- `for (i in -3:7) {print(i^2)}`→ 印出index-3~7的平方

```
> for (i in -3:7) {print(i^2)}
[1] 9
[1] 4
[1] 1
[1] 0
[1] 1
[1] 4
[1] 9
[1] 16
[1] 25
[1] 36
[1] 49
```

- `for (i in -3:1) {for (j in 7:9) {print(i^2+sqrt(j))}}` → `i`先以-3帶入跑j=6-9的迴圈，共跑四次，每個*i*會個跑一次j=6-9的迴圈

```
> for (i in -3:1) {for (j in 7:9) {print(i^2+sqrt(j))}}
[1] 11.64575
[1] 11.82843
[1] 12
[1] 6.645751
[1] 6.828427
[1] 7
[1] 3.645751
[1] 3.828427
[1] 4
[1] 2.645751
[1] 2.828427
[1] 3
[1] 3.645751
[1] 3.828427
[1] 4
```

- `apply(X=data1, MARGIN=2, MEAN)`，在資料data1中以行為單位算出每行平均值；
MARGIN=1表示適用於每一列，MARGIN=2表示適用於每一行；
MEAN可以替換成其他想要套用的功能，如畫折線圖(`X=data1, MARGIN=2, FUN=plot, type="l"`)
- `tapply(X, INDEX, FUN)`，用INDEX分類X的資料
 - `tapply(age, smoke, mean)`，分別算出吸煙者和不吸菸者的平均年齡